Probing a Model of a GPCR/Ligand Complex in an Explicit Membrane Environment: The Human Cholecystokinin-1 Receptor

Jérôme Hénin,* Bernard Maigret,* Mounir Tarek,* Chantal Escrieut,† Daniel Fourmy,† and Christophe Chipot*
*Equipe de Dynamique des Assemblages Membranaires, UMR CNRS/UHP 7565, Institut Nancéien de Chimie Moléculaire, Université Henri Poincaré, Vandœuvre-lès-Nancy, France; and †Equipe de Biologie et Pathologie Digestive, INSERM U531, Institut Louis Bugnard, Toulouse, France

ABSTRACT A three-dimensional model structure of a complex formed by a G-protein-coupled receptor (GPCR) and an agonist ligand is probed and refined using molecular-dynamics simulations and free energy calculations in a realistic environment. The model of the human receptor of cholecystokinin associated to agonist ligand CCK9 was obtained from a synergistic procedure combining site-directed mutagenesis experiments and in silico modeling. The 31-ns molecular-dynamics simulation in an explicit membrane environment indicates that both the structure of the receptor and its interactions with the ligand are robust. Whereas the secondary structure of the α-helix bundle is well preserved, the region of the intracellular loops exhibits a significant flexibility likely to be ascribed to the absence of G-protein subunits in the model. New insight into the structural features of the binding pocket is gained, in particular, the interplay of the ligand with both the receptor and internal water molecules. Water-mediated interactions are shown to participate in the binding, hence, suggesting additional site-directed mutagenesis experiments. Accurate free energy calculations on mutated ligands provide differences in the receptor-ligand binding affinity, thus offering a direct, quantitative comparison to experiment. We propose that this detailed consistency-checking procedure be used as a routine refinement step of in vacuo GPCR models, before further investigation and application to structure-based drug design.

INTRODUCTION

One of the grand challenges amenable to molecular modeling is to provide a microscopic insight into membrane proteins when conventional experimental techniques are not able to supply this level of information. Of particular interest are seven transmembrane (TM) domain G-protein-coupled receptors (GPCRs) (1), which correspond to the third largest family of genes in the human genome (2,3), and constitute key targets for de novo drug design (4). Alone, molecular modeling is unlikely to reach the atomic detail hitherto inaccessible to experiment, but is anticipated to help interpreting inferences accrued in recent years from the wealth of structure-activity data available for GPCRs, thereby improving our understanding of their ontogeny and function (5). Resolution of the three-dimensional structure of bovine rhodopsin (6), the only high-resolution GPCR structure determined by x-ray crystallography to this date, has opened new prospects for the modeling of structurally related membrane proteins, by providing valuable guidelines against which original theoretical hypotheses can be confronted (7,8). Unfortunately, these guidelines suffer severe limitations rooted in the conformation in which rhodopsin was crystallized, namely, the inactive, dark state of the receptor, recognized in some instances to be an unsuitable template for activated complexes formed by a GPCR and an agonist ligand (9–11).

In the past five years, much effort has been invested in the prediction of the structure and function of GPCRs, employing molecular modeling tools ranging from first principles (12,13) to knowledge-based methods (7,14,15). In the first class of approaches, the individually predicted and optimized TM helical segments of the modeled GPCR are packed using a template of known structure, usually rhodopsin or, in earlier investigations, bacteriorhodopsin. The second class of approaches corresponds to a homology modeling relying upon the crystallographic structure of a reasonably related receptor, together with available experimental information. In practice, such schemes are restricted to the rhodopsin template and, consequently, suffer from its inherent limitations when targeting activated receptors. An alternative to the latter methods consists of a fully physics-based conformational search using well-identified TM sequences, thereby obviating the need for a template of known structure (16).

Molecular constructs resulting from knowledge-based approaches can be employed advantageously to design site-directed mutagenesis experiments, which, in turn, can serve to refine the models and understand the function of the receptor. A recent contribution (17) combines experimental and theoretical approaches in a self-consistent fashion, to model the human receptor of cholecystokinin (CCK) (18). CCK is ubiquitous to the gastrointestinal and central nervous systems, where it acts as a hormone and a neurotransmitter, respectively. Control of satiety, gallbladder contraction, pancreatic exocrine secretions, gastric emptying, and gut motility constitute pivotal actions of CCK that are mediated by the so-called CCK1 receptor (CCK1R). The modeled complex is formed by CCK1R and an agonist ligand (17), the nonapeptide (Met, Nle)-CCK9, an analog of the endogenous

Submitted July 12, 2005, and accepted for publication October 24, 2005.

Address reprint requests to Christophe Chipot, Tel.: 33-3-83-68-40-97; E-mail: christophe.chipot@edam.uhp-nancy.fr.

0006-3495/06/02/1232/09 $2.00

doi: 10.1529/biophysj.105.070599

ligand with the sequence Arg-Asp-S-Tyr-Thr-Gly-Trp-Met-Asp-Phe-NH_2—referred to as CCK9 in what follows. The effort that ultimately yielded this model structure involved a synergy between experiments and modeling (11,19–24). Two-dimensional mutagenesis experiments guided the refinement of the model, which in turn highlighted possible interactions and led to the design of additional mutants, thereby offering a cross-validation and complementarity between experimental and theoretical information.

Despite the constraining framework of experimental data, the proposed structure remains an in vacuo model, which does not offer any guarantee of consistency with the physical conditions of a membrane environment. Furthermore, interpretation of binding-affinity assays in terms of well-localized interactions, only constitute an indirect link between the model and experimental evidence. The route chosen here to probe the behavior of the CCK1R model consists in performing molecular dynamics and free energy simulations in a realistic hydrated lipid bilayer. Achieving a good level of consistency between simulated and measured properties should increase the degree of confidence in the structure when it is used in docking studies aimed at screening potentially active ligands. Noteworthily, screening alone has been applied to appraise the reliability of GPCR models (7,25,26). Docking procedures have also been used to predict binding modes in GPCR models and estimate receptor-ligand affinities (12). The strong dependence of molecular docking to the scoring functions, however, makes it less reliable than free-energy calculations. The significant computational effort involved in free energy calculations explains why, hitherto, the latter have not been applied to theoretical models of receptor-ligand assemblies, in particular to those only partially validated. Molecular dynamics (MD) simulations in an explicit water-membrane environment have, however, been employed in a handful of instances to probe the structure of GPCRs, either the prototypical rhodopsin (27), or models of not-yet-resolved three-dimensional structures, like the CXCR4 (28), the μ-opioid (29), or the δ-opioid (30) receptor.

In this article, large-scale molecular dynamics simulations are used to probe and refine the model of CCK1R:CCK9 when embedded in a lipid bilayer, with particular attention to the conservation of the experimental constraints. In addition, this step provides an insight into the interplay of the receptor-ligand complex with the explicit molecules forming its environment, which are clearly absent from the in vacuo construct. Next, a direct, quantitative criterion for comparison to experiment is obtained through free energy calculations targeted at reproducing relative binding affinities measured for mutated ligands. Such alchemical transformations are employed for the first time to close-the-loop of the modeling process of a GPCR, by reproducing in silico the experimental binding affinities that were utilized to guide its construction. The point mutations chosen here correspond to the replacement of sulfated tyrosine S-Tyr-3 by a tyrosyl residue and of Asp-8 by alanine—namely, the third and penultimate amino acids at the N- and the C-termini of CCK9, respectively interacting with the extracellular loop and buried in the binding pocket. Selection of these transformations was dictated by the limited magnitude of the structural changes in the agonist ligand and, a priori, in the receptor, and the necessity of marked differences in the binding affinities, compatible with the level of accuracy currently attained by free energy calculations (31). Alanine replacement of Asp-8 at the C-terminus of CCK9 and removal of the sulfate moiety of S-Tyr-3 at the N-terminus have been demonstrated to yield a significant decrease in the binding affinity of CCK9 toward CCK1R (20–23). Furthermore, the two mutated ligands retain the biological activity of the agonist (23). Mutations in the bound state are, therefore, not expected to entail an allosteric transition of the receptor, which would occur over timescales exceeding those accessible to all-atom MD simulations.

METHODS

Description of the system

Considering that the first residues of CCK1R do not play any particular role on binding affinities (19), it was chosen to truncate the native sequence of 376 amino acids at its N-terminus, resulting in a model of the receptor that consists of 345 residues. The receptor-ligand complex was embedded in a fully hydrated palmitoyloleoylphosphatidylcholine (POPC) bilayer. Initially, the receptor-ligand complex was positioned across the equilibrated bilayer, while seeking to match the hydrophobic protein segments with the layer formed by the lipid hydrocarbon tails. Lipids overlapping with the protein complex were deleted, leaving a bilayer consisting of 228 POPC molecules. To ascertain that the cytoplasmic and extracellular loops do not interact, an amount of 16,527 water molecules was added, as well as 24 chloride counterions to compensate for the positive net charge of the protein, thus making a total number of atoms equal to 72,255. The complete system, represented in Fig. 1, was replicated periodically in the three directions of space, with a vertical repeat distance of $\sim$115 Å. For the simulation of the ligand in aqueous medium, the solvent box contained 4120 water molecules, one chloride and two sodium counterions.

Molecular dynamics simulations

A 31-ns MD simulation was carried out in the isobaric-isothermal ensemble, maintaining the pressure and the temperature at 1.0 atm and 300.0 K, respectively, by means of Langevin dynamics and the Langevin piston approach. The MD program NAMD (32,33) was employed in conjunction with the CHARMM27 force field (34) to describe the receptor, the agonist ligand, the lipid bilayer, and the water molecules. Since the sulfated tyrosine residue is not described in the standard force field, new dihedral angle parameters were computed and a set of point charges was derived from the molecular electrostatic potential. This parameterization step was based on ab initio calculations at the RHF/6-31G* level. A united-atom description of the alkyl chains of the lipid molecules was utilized. Coulomb forces were evaluated with the particle-mesh Ewald method. The equations of motion were integrated with a 2-fs time step, using the r-RESPA algorithm to update short- and long-range contributions at different frequencies.

Free energy calculations of alchemical transformations

Direct insertion of the agonist ligand in the binding site of the protein by means of a potential of mean-force-like calculation, which would yield an

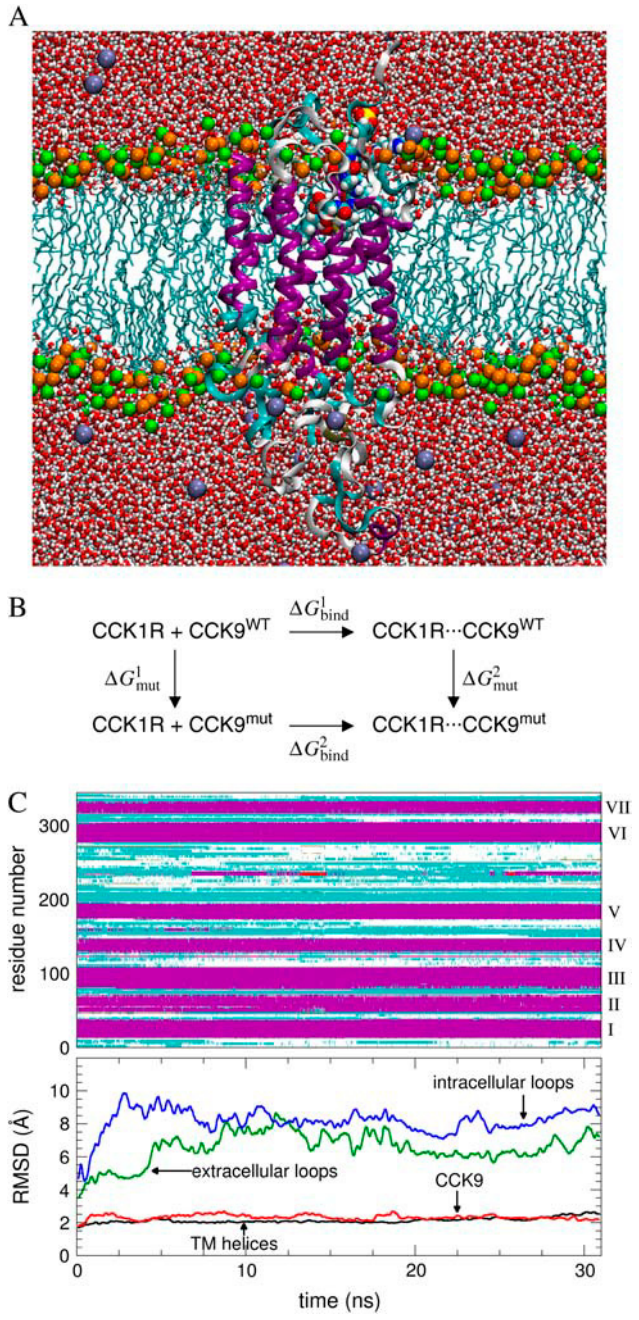

FIGURE 1 (*A*) Front view of the assembly formed by CCK1R and CCK9 in an hydrated POPC bilayer. The α- and 3_{10}-helices are represented as purple and pink ribbons, respectively. Coils and β-turns are shown as white and cyan tubes. CCK9 is depicted as van der Waals spheres. P- and N-atoms of the POPC headgroups are highlighted as orange and green spheres, respectively. Atoms of water molecules are shown as van der Waals spheres. Lipid chains are represented as cyan rods. Image rendering was obtained with VMD (53). (*B*) Thermodynamic cycle used for the alchemical transformations of the wild-type (WT) ligand CCK9 in the free, hydrated state (*left*) and in the bound state (*right*). The values ΔG^1_{bind} and ΔG^2_{bind} are determined experimentally. The free energy difference between the WT and mutant ligands is given by $\Delta G^2_{\text{bind}} - \Delta G^1_{\text{bind}} = \Delta G^2_{\text{mut}} - \Delta G^1_{\text{mut}}$. (*C*) Secondary structure of CCK1R as a function of time (*top*). The color scheme for α- and 3_{10}-helices, β-turns, and coils follows the same conventions as *a*. Roman numerals on the right-hand side indicate the TM helix number. Distance RMSD (α-carbon atoms only) of various components of the system with respect to the initial structure, as a function of time (*bottom*). The time axis is common to both graphs.

absolute free energy of association, still remains out of reach for current MD simulations, on account of the slow relaxation of the collective degrees of freedom in the membrane protein and its surroundings. To circumvent this difficulty, for both the Asp-8 to Ala and the S-Tyr-3 to Tyr transformations, free energy differences were computed between two distinct ligands—the wild-type (CCK9), and the mutant—both in the free state, i.e., in an aqueous environment, and bound to the receptor, as depicted in the thermodynamic cycle of Fig. 1. Point mutations in bulk water and in the receptor were performed employing the free energy perturbation (FEP) method (35), wherein the Gibbs free energy difference between two thermodynamic states connected by M intermediate, nonphysical substates is expressed as

$$\Delta G = -\frac{1}{\beta}\sum_{i=1}^{M+1} \ln\langle \exp\{-\beta[\mathcal{V}(\mathbf{x};\lambda_{i+1}) - \mathcal{V}(\mathbf{x};\lambda_i)]\}\rangle_{\lambda_i},$$
$$\beta = 1/k_B T,$$

where k_B is the Boltzmann constant, T, the temperature; and $\mathcal{V}(\mathbf{x};\lambda_i)$, the potential energy function that depends upon the Cartesian coordinates of the system $\{\mathbf{x}\}$, and the coupling parameter, λ_i, that connects the initial and the final states of the transformation. Considering the nature of the point mutations, namely, neutralization of a charged amino acid, it was chosen to break the reaction path into 114 stages of uneven widths. Narrow intermediate states were defined toward the end points of the simulation to avoid singularities due to ignited interactions of an appearing moiety with its environment. Each transformation, either in bulk water or in CCK1R (see Fig. 1), was run for 3.4 ns. The dual-topology paradigm was utilized, wherein the initial and the final states are defined concomitantly, but do not interact (36). Simulations of CCK9 in water included a small number of explicit counterions, which may lead to severe convergence issues in free energy calculations (37). This was alleviated by enforcing harmonic positional restraints on the sodium and chloride ions. Alchemical transformations involved both a negatively charged side chain and a sodium counterion, so that the overall charge of the system was zero throughout the transformation.

Estimation of errors in free energy calculations is usually difficult and presupposes stringent underlying approximations. It has been chosen to provide an estimate of the error based on 1), carrying out the same alchemical transformations twice, after 5.0 and 10.5 ns of MD trajectory; and 2), a first-order expansion of the free energy,

$$\Delta G = -\frac{1}{\beta}\ln\{\langle \exp[-\beta\Delta\mathcal{V}(\mathbf{x};\lambda)]\rangle_\lambda \pm \delta\varepsilon\},$$

where $\delta\varepsilon$ is the statistical error on the ensemble average, $\langle \exp[-\beta\Delta\mathcal{V}(\mathbf{x};\lambda)]\rangle_\lambda$, expressed as

$$\delta\varepsilon^2 = \frac{1+2\tau}{N}\{\langle \exp[-2\beta\Delta\mathcal{V}(\mathbf{x};\lambda)]\rangle_\lambda - \langle \exp[-\beta\Delta\mathcal{V}(\mathbf{x};\lambda)]\rangle_\lambda^2\}.$$

Here, N is the number of samples accrued in the FEP calculation, and $(1 + 2\tau)$ is the sampling ratio of the latter (38). Assuming that $\delta\varepsilon$ is appreciably smaller than $\langle \exp[-\beta\Delta\mathcal{V}(\mathbf{x};\lambda)]\rangle_\lambda$, the free energy change can be written at the first-order as

$$\Delta G = -\frac{1}{\beta}\left\{\ln\langle \exp[-\beta\Delta\mathcal{V}(\mathbf{x};\lambda)]\rangle_\lambda \pm \frac{\delta\varepsilon}{\langle \exp[-\beta\Delta\mathcal{V}(\mathbf{x};\lambda)]\rangle_\lambda}\right\}.$$

Binding-affinity assays

New affinity measurements of CCK9 analogs toward CCK1R were carried out in order to provide refined binding free energies with an increased statistical confidence over the ones published previously (21,22). The binding affinities were estimated by measuring in vitro the radioactivity of transfected cells expressing the receptor, after exposure to radio-labeled ligands. Nonapeptides were synthesized as described in previous reports (39,40). ^{125}INa was obtained from Pharmacia-Amersham (Les Ulis, France). The CCK9 analogs were conjugated with the Bolton-Hunter reagent, purified and radioiodinated, as described in Fourmy et al. (41). The specific activity of the radioiodinated peptide was 1600–2000 Ci/mmol. All other chemicals were obtained from commercial sources. COS-7 cells (1.5×106) were plated onto 10-cm culture dishes and grown in Dulbecco's modified Eagle's medium containing 5% of fetal calf serum in a 5% CO_2 atmosphere, at 37°C. After overnight incubation, cells were transfected with 2.5 g/plate of pRFENeo vectors containing the cDNA for the wild-type (WT) receptor, using a modified DEAE-dextran method. Cells were then transferred onto 24-well plates at a density of 80,000 to 150,000 cells/well, 24 h after transfection. Approximately 24 hours after the transfer, the cells were washed with pH 6.95, 0.1% BSA phosphate buffer and then incubated for 60 min at 37°C in 0.5 ml of Dulbecco's modified Eagle's medium, 0.1% BSA with either 71 pM 125I-BH-(Thr, Nle)-CCK9 in the presence or the absence of competing compounds. Nonspecific binding was determined in the presence of 1 μM ligand. The cells were washed twice with cold, 2% BSA phosphate buffer, and the cell-associated radioligand was collected with 0.1 M NaOH added to each well. The radioactivity was directly counted using a γ-counter.

RESULTS AND DISCUSSION

Molecular dynamics simulation in a membrane environment

To probe the intrinsic stability and the dynamic behavior of the assembly built in vacuo and formed by CCK1R and the agonist ligand CCK9 (17), the complete model was inserted in a POPC bilayer and a 31-ns MD trajectory was generated. The slow rearrangement of the aliphatic chains around the receptor, occurring on the nanosecond timescale, is reflected in a moderate contraction of the simulation cell in the plane of the lipid membrane. Beyond ~5 ns, fluctuations of the dimensions of the system and in its total energy are sufficiently limited to assume safely that an equilibrium has been reached. Short, subnanosecond MD simulations that are incompatible with the timescale of the organization of lipid molecules near the receptor, should, therefore, be interpreted with particular care.

The distance root mean-square deviation (RMSD) shown in Fig. 1, computed using the backbone atoms of the integral α-helices, with respect to the in vacuo construct, remains close to 2.0 Å over the first 20 ns, before increasing to ~2.5 Å in the remaining 11 ns, thereby highlighting the slow relaxation of the collective degrees of freedom in the TM domain. Deconvolution of the RMSD reveals an initial greater disorder on the cytoplasmic side of the membrane, the extracellular loops being partially structured by the ligand protruding from the binding site. Considering that the positioning of the intra- and extracellular loops has been performed in vacuo, based on a limited amount of experimental data, and, more importantly, in the absence of the G-protein subunits, significant rearrangement is to be expected when immersed in a solvent. As has been underlined previously, failure to take into account the environment of the loop region in an appropriate fashion may entail severe distortions in the structure of the receptor (42). A closer inspection of the MD trajectory reveals that, on account of the structural fluctuations in the intracellular loops, the cytoplasmic end of the α-helices is prone to depart from the initial construct. In sharp contrast, the remainder of the TM domain is well anchored around the binding site and does not undergo noticeable structural modifications.

Secondary structure analysis of the complete protein (Fig. 1) indicates that the α-helical motifs in the TM region are well conserved throughout the trajectory. Though flexible, the core of CCK9 is found predominantly in a β-turn conformation (43), in line with solution NMR analyses, which suggest a dynamic equilibrium between ordered structures (44,45).

Stability of the α-helix bundle over the simulated timescale indicates that the system has evolved around the same local energy minimum. One should note, however, that, seemingly stable structures in the gas phase are not necessarily local energy minima of the molecular assembly formed by the protein inserted in a water-lipid bilayer arrangement. Besides, speaking of stability in vacuo is often equivocal, considering that in the absence of appropriately chosen constraints, disruption of the tertiary structure may rapidly occur.

The essential of the key protein-ligand interactions brought to light experimentally are preserved throughout the simulation (see Fig. 2 *A*). Of particular interest, a network of hydrophobic residues of CCK1R, that includes Leu-53, Val-125, Ile-329, and Ile-352, appears to participate in favorable van der Waals contacts with amino acids Met-7 and Phe-9 of the agonist ligand (24). In addition, the strong electrostatic interactions of Arg-336 with Asp-8 (22), and of Met-195 and Arg-197 with the sulfated tyrosine S-Tyr-3, are equally well conserved (21,23). Photoaffinity labeling experiments also confirm the direct interaction of Arg-197 with the sulfated tyrosyl residue of the agonist ligand (46). Most importantly, NMR investigations of CCK interacting with a fragment of CCK1R that encompasses the top region of TM helix VI and the third extracellular loop further supports the present docking mode (44), hence casting doubt on the conjecture of a reverted ligand that would interact with Trp-39 at its C-terminus (43,47,48).

Mobility of the N-terminal segment of the receptor causes a loss of contacts between Trp-39 and Gln-40 with the N-terminus of CCK9 (20), substituted by a hydration shell. It is legitimate to wonder whether this loss of interactions results from the low specificity of the latter, or is a reflection of a local flaw in the model. Site-directed mutagenesis of amino acids Trp-39 and Gln-40 leads to an increase in the dissociation constant, K_d, <20-fold (20). For comparison purposes, replacement of Arg-336 by methionine corresponds

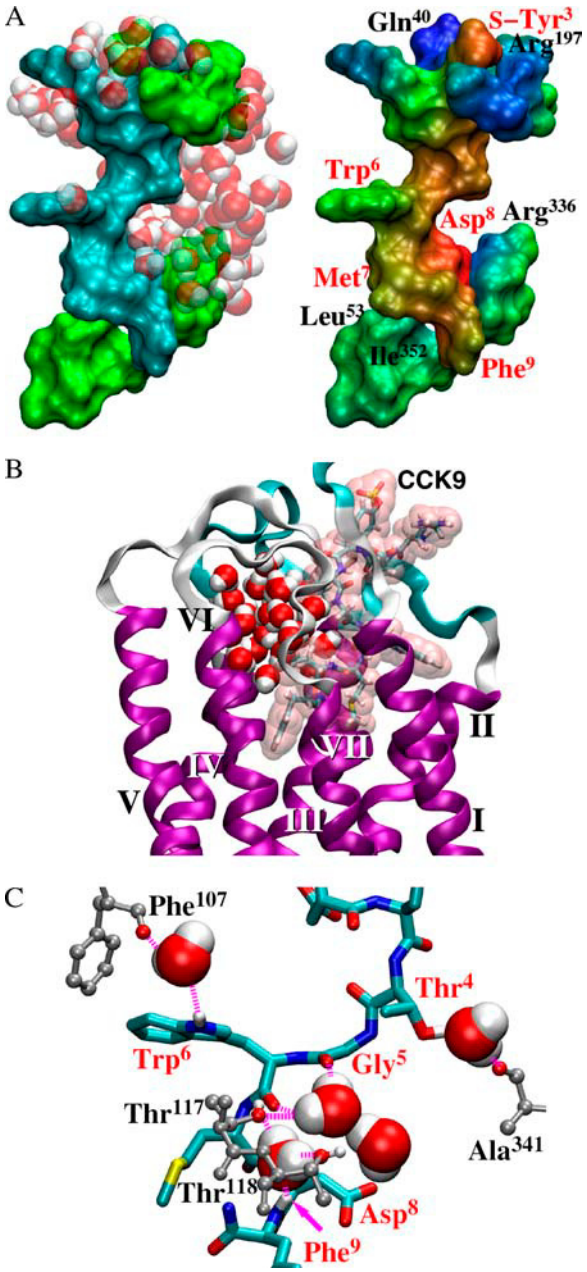

FIGURE 2 (*A*) Insight into the environment of the agonist ligand (*left*). CCK9 is represented as a cyan surface, while interacting residues of CCK1R brought to light experimentally are shown as green surfaces. Surrounding water molecules are displayed as semitransparent spheres. Electrostatic complementarity of the ligand and its receptor (*right*). Surfaces are colored according to molecular electrostatic potentials. (*B*) Hydration pattern in the binding site of CCK1R. CCK9 is depicted in a transparent van der Waals envelope. (*C*) Indirect, water-mediated interactions of CCK9 with CCK1R. Residues of the receptor are shown in a gray ball-and-stick representation. Hydrogen bonds are highlighted as magenta dashed lines. The magenta arrow points to the hydrogen bond formed between the amino group of Phe-9 and a bridging water that interacts concomitantly with Thr-117 and Thr-118.

to an increase of K_d by a factor ~118, thus suggesting that receptor-ligand interactions in the N-terminal region of CCK1R play an appreciably lesser role in the binding of CCK9 (22).

The use of an explicit bilayer environment may also be viewed as an enhancement to the model, as it grants access to a detailed view of molecular interactions involving solvent molecules. The binding pocket of CCK1R is sufficiently wide to allow water molecules from the bulk to flow in and interact with the agonist ligand, which is consistent with previous observations in bacteriorhodopsin (49). Whereas hydration of CCK9 is substantial at its N-terminus, in the region of the extracellular loops, water molecules become scarce as the peptide is buried deeper in the receptor and are confined in a well-localized pocket that contains an average of 25 molecules and communicates with the bulk (see Fig. 2 *B*). This water cluster, obviously absent in the original gas phase model, progressively forms in the cavity over a period of 3 ns, concomitantly with the creation of a passageway between TM helices III and IV. Residence times for molecules near the C-terminus of CCK9 may exceed the nanosecond timescale, in sharp contrast with those located higher in the crevice. Near the extracellular loops, water molecules follow a diffusive regime and exchange rapidly with the bulk aqueous medium. It should be noted that internal water molecules do not disrupt the key receptor-ligand interactions brought to light experimentally, but instead, fill available vacuities and solvate moieties of the ligand that are not interacting directly with CCK1R (see Fig. 2 *A*).

The direct receptor-ligand interactions, which constitute the scaffold of the CCK1R:CCK9 association, are only elements of the complete binding mode. Several indirect interactions, relayed by water molecules, also appear to participate in the binding. At the C-terminal of CCK9, the amino group of Phe-9 interacts with the neighboring Thr-117 and Thr-118 residues through a single bridging water molecule (see Fig. 2 *C*). Further up, the indole −NH group of Trp-6 interacts with Phe-107 by means of a bridged water molecule, and the carbonyl group of Gly-5 interacts with the hydroxyl moiety of Thr-117 relayed by a water molecule. A water molecule pertaining to the aqueous pocket depicted in Fig. 2 *B* connects Thr-4 to Ala-341. Moving outside the crevice, water-mediated interactions of the remainder of the agonist ligand with CCK1R are shorter lived, due to the faster exchange of the bridging water molecules with the bulk aqueous medium.

The original binding mode is virtually unperturbed as water molecules propagate along in the crevice—a result that is likely to stem from the strong anchoring of CCK9 at its C-terminus, namely, Asp-8, and at its N-terminus, S-Tyr-3, reinforced by steady van der Waals contacts formed between both Met-7 and Phe-9 and the hydrophobic pocket. Alternative binding modes for other agonist or antagonist ligands, and how water molecules mediate their association to

CCK1R should be considered in the light of the present structural data. Of particular interest, the role played by the vicinal Thr-117 and Thr-118 residues in the water-assisted binding of ligands deserves further attention and should be probed through site-directed mutagenesis experiments.

Receptor-ligand binding free energies

In terms of relation to experiment, the present MD simulations essentially indicate whether or not the key protein-ligand interactions are conserved through time. Free energy calculations go one step beyond, by quantifying the importance of these interactions, and, hence, represent a tangible thermodynamic measure for appraising the reliability of the construct. Experimental and computed relative binding free energies for the Asp-8 to Ala and S-Tyr-3 to Tyr point mutations in CCK9 are gathered in Table 1.

The FEP estimate of $+3.0 \pm 0.7$ kcal/mol for the Asp-8 to Ala transformation agrees remarkably well with the binding-affinity assays experiments repeated up to six times, yielding a free energy change of $+3.2 \pm 0.3$ kcal/mol. At the microscopic level, replacement of Asp-8 by Ala is accompanied by subtle modifications in the crevice of CCK1R. Around the middle of the mutation, i.e., $\lambda = 0.5$, Arg-336 pivots to interact with residue Thr-4 of CCK9. Interaction of Asp-8 with Asn-333 is, however, preserved. Toward the end of the transformation, the guanidinium moiety of Arg-336 forms a hydrogen bond with the amide group of Asn-333, but the corresponding conformational modifications of the side chains appear to be sufficiently limited to not perturb the overall arrangement of the TM helices.

In silico replacement of S-Tyr-3 by Tyr led to an estimate of $+1.9 \pm 0.4$ kcal/mol based on two different initial sets of coordinates and momenta, whereas experiment predicts a free energy change of $+2.7 \pm 0.1$ kcal/mol. During this point mutation, binding of the tyrosyl residue to Arg-197 progressively weakens, as the level of hydration of the latter increases. Though more flexible than the lower end of the binding pocket, the structure of the extracellular loops remains essentially unperturbed by the alchemical transformation. This flexibility, which can only be fully captured over long timescales, is envisioned to affect the accuracy of the simulation, thereby explaining the imperfect accord between the theoretical and the experimental estimates of the binding free energy difference, not necessarily reflected in the statistical precision of the former. Reaching the desired level of accuracy implies an appropriate description of slowly relaxing phenomena, which, in most cases, constitutes a major obstacle in free energy calculations of large, biologically relevant systems, like membrane receptors. Another possible source of error can be found in the suboptimal parameterization of the nonstandard sulfated tyrosine residue.

Analysis of the convergence properties of the present simulations indicates a smooth behavior of the free energy as a function of λ and converging ensemble averages for the different λ-intermediates between the initial and the final states of the alchemical transformation (see Fig. 3). Furthermore, configurational ensembles characteristic of contiguous states appear to overlap very well, thereby satisfying a necessary condition for appropriate convergence of FEP calculations (50).

Interestingly enough, compared to the FEP approach, application of the acceptance ratio method of Bennett (50,51) led to an agreement within 0.2 kcal/mol for the Asp-8 to Ala mutation, thereby suggesting appropriate convergence of the FEP ensemble averages. In the case of the S-Tyr-3 to Tyr transformation, the differences in ΔG^1_{mut} and ΔG^2_{mut} are also within 0.2 kcal/mol, but in opposite directions, thus yielding a net binding free energy of $+2.2$ kcal/mol, somewhat closer to the target experimental value than the FEP estimate. It is worth underlining that for both point mutations, the difference in the binding free energies between FEP and Bennett's approach is always less than the error estimate of the calculation, which would imply that convergence properties of the method is not a critical issue here. In contrast with the error estimate reached from the simulations at 5.0 and 10.5 ns, the first-order expansion only reflects the statistical precision of the computation—namely, on the order of ± 0.3 kcal/mol, which clearly does not account for fluctuations in the receptor-ligand structure over long timescales. It would, therefore, appear that error estimates based on the statistical properties of a single free energy calculation can be deceiving, as they only provide the fraction of the total error corresponding to the sampled region of the configurational space. Arguably, relaxation along the slowest manifolds cannot be captured in the ensemble averages, but comparison of the latter in distinct simulations is expected to provide an appraisal of the accuracy of the free energy calculations.

TABLE 1 Experimental and calculated relative binding free energies for the Asp-8 to Ala and S-Tyr-3 to Tyr point mutations in agonist ligand CCK9

	Experiment			Theory		
Point mutation	K_i (WT)	K_i (mut)	ΔG_{bind}	ΔG^1_{mut}	ΔG^2_{mut}	ΔG_{bind}
Asp-8 → Ala	1.38 ± 0.15	253.8 ± 11.4	$+3.2 \pm 0.3$	$+226.2$	$+229.2$	$+3.0 \pm 0.7$
S-Tyr-3 → Tyr	1.38 ± 0.15	108.8 ± 4.8	$+2.7 \pm 0.1$	$+181.1$	$+183.0$	$+1.9 \pm 0.4$

All experimental binding constants are in nanomoles, and free energies in kcal/mol. The notations utilized here are described in the thermodynamic cycle of Fig. 2. Error estimates are obtained by repeating the free energy calculations with two different starting points, at 5 and 10.5 ns of the MD trajectory.

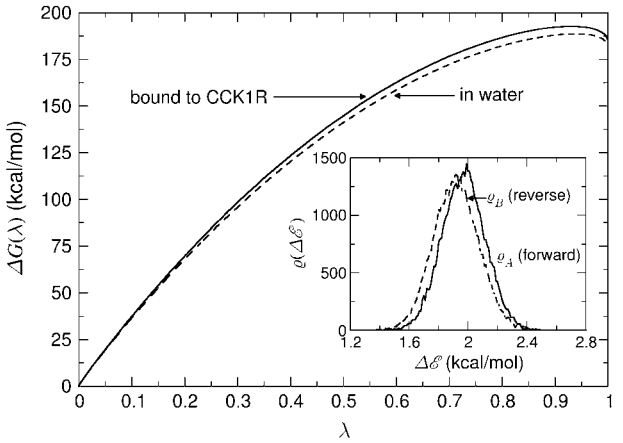

FIGURE 3 Evolution of the Gibbs free energy as a function of the coupling parameter λ. Here, the replacement of the S-Tyr-3 by tyrosine in CCK9 is shown in the free and in the bound states, starting from a configuration obtained at 10.5 ns of the molecular dynamics (MD) trajectory. (*Inset*) Typical overlap of configurational ensembles reflected in their respective density of states, $\varrho(\mathcal{E})$, characteristic of contiguous intermediate states of the transformation, at $\lambda = 0.5$.

CONCLUSIONS

The paucity of high-resolution structural information for membrane proteins and, in particular, GPCRs, has opened new vistas for molecular modeling. Among the methods that have been devised to model GPCRs, the synergistic combination of theory and experiment constitutes a promising perspective. There is no guarantee, however, that a model built in vacuo will behave similarly in a more relevant, heterogeneous water-lipid medium (12).

In this contribution, the behavior of a three-dimensional model of a GPCR (17) in an explicit membrane environment has been probed by means of large-scale MD simulations. Stability over 31 ns of the receptor-ligand complex embedded in a POPC bilayer suggests that the system lies in a local minimum of the configurational space, which satisfies the experimental restraints used to build the model. All-atom simulations in realistic surroundings can be viewed as a refinement of the underlying physical criteria upon which the construct is based, while preserving consistency with the body of experimental information. Whereas interpretation of binding-affinity measurements in terms of key interactions constitutes an indirect link between experiment and theory, free energy calculations open the way to a quantitative comparison of biologically relevant properties of the system. The very encouraging reproduction of relative binding free energies at the N- and the C-termini of agonist ligand CCK9 supports the view that the model complies with the experimental observations employed in the process of its construction (17). Enhancement of the model by means of an explicit membrane environment offers new insight into the role played by the solvent on the receptor-ligand complex. Of particular interest, the diffusion of water molecules in the binding pocket of CCK1R is not incompatible with the binding mode predicted in the in vacuo construct. It further highlights amino acids of the receptor with which the ligand interacts indirectly by means of discrete water molecules. Confirming the role played by these residues in the formation of the receptor-ligand complex to further understand how the latter is mediated by internal water molecules is expected to prompt new site-directed mutagenesis experiments. In turn, improved characterization of the binding mode should provide a valuable help in the design of potent agonist and antagonist ligands.

Put together, the present results have increased our level of confidence in the original model, restrained to the properties of the binding pocket and the TM α-helix bundle, thereby paving the way to new hypotheses for interpreting available structure-activity data. In this sense, it should be advocated that the present molecular assembly, including the environment, be used for further studies, e.g., molecular docking of alternate ligands, in particular in the framework of virtual screening. To this effect, for completeness, the Cartesian atomic coordinates of the receptor-ligand complex and its internal water molecules are provided as Supplemental Material. It still remains that the level of reliability in the model of CCK1R is largely uneven, and markedly weaker in the case of the intracellular loops, mostly due to the absence of the G-protein subunits in the modeling of the latter.

Nevertheless, not only does the model reflect the experimental observations utilized for its construction, but, more importantly, it is capable of reproducing accurately thermodynamic quantities that were not fed explicitly into the modeling process. Among all the models of GPCRs published hitherto, how many would actually fulfill these criteria? It would be desirable that the simulation protocols described herein be employed routinely for testing in vacuo models before any further use.

The present strategy, however, is not expected to supply a model of a GPCR closely matching its native three-dimensional structure. Besides, it is far from evident to assess how similar a model and a functional receptor are, inasmuch as the structure has not been resolved experimentally. Awaiting the release of high-resolution data, carefully devised and tested models are anticipated to continue integrating new information, and provide a reliable structural basis (52) that can be utilized to advance our understanding of the function of GPCRs and discover novel therapeutic agents.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

We gratefully acknowledge the Centre Charles Hermite, Vandœuvre-lès-Nancy, France, and the Centre Informatique National de l'enseignement Supérieur (CINES), Montpellier, France, for provision of generous amounts of computational time on their SGI Origin 3000.

We are indebted to the Association pour la Recherche sur le Cancer for research grant No. 3282.

REFERENCES

1. Bockaert, J., and J. P. Pin. 1999. Molecular tinkering of G-protein-coupled receptors: an evolutionary success. *EMBO J.* 18:1723–1728.
2. Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, et al. 2001. The sequence of the human genome. *Science.* 291:1304–1351.
3. Takeda, S., S. Kadowaki, T. Haga, H. Takaesu, and S. Mitaku. 2002. Identification of G-protein-coupled receptor genes from the human genome sequence. *FEBS Lett.* 520:97–101.
4. Wise, A., K. Gearing, and S. Rees. 2002. Target validation of G-protein coupled receptors. *Drug Discov. Today.* 7:235–246.
5. Visiers, I., J. A. Ballesteros, and H. Weinstein. 2002. Three-dimensional representations of G-protein-coupled receptor structures and mechanisms. *Methods Enzymol.* 343:329–371.
6. Palczewski, K., T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, and M. Miyano. 2000. Crystal structure of rhodopsin: a G-protein-coupled receptor. *Science.* 289:739–745.
7. Bissantz, C., P. Bernard, M. Hibert, and D. Rognan. 2003. Protein-based virtual screening of chemical databases. II. Are homology models of G-protein coupled receptors suitable targets? *Proteins.* 50: 5–25.
8. Gouldson, P. R., N. J. Kidley, R. P. Bywater, G. Psaroudakis, H. D. Brooks, C. Diaz, D. Shire, and C. A. Reynolds. 2004. Toward the active conformations of rhodopsin and the β_2-adrenergic receptor. *Proteins.* 56:67–84.
9. Gershengorn, M. C., and R. Osman. 2001. Insights into G-protein-coupled receptor function using molecular models. *Endocrinology.* 142:2–10.
10. Manivet, P., B. Schneider, J. C. Smith, D. S. Choi, L. Maroteaux, O. Kellermann, and J. M. Launay. 2002. The serotonin binding site of human and murine 5-HT2B receptors. *J. Biol. Chem.* 19:17170–17178.
11. Archer, E., B. Maigret, C. Escrieut, L. Pradayrol, and D. Fourmy. 2003. Rhodopsin crystal: new template yielding realistic models of G-protein-coupled receptors? *Trends Pharmacol. Sci.* 24:36–40.
12. Vaidehi, N., W. B. Floriano, R. Trabanino, S. E. Hall, P. Freddolino, E. J. Choi, G. Zamanakos, and W. A. Goddard III. 2002. Prediction of structure and function of G-protein-coupled receptors. *Proc. Natl. Acad. Sci. USA.* 99:12622–12627.
13. Becker, O. M., S. Shacham, Y. Marantz, and S. Noiman. 2003. Modeling the 3D structure of GPCRs: advances and application to drug discovery. *Curr. Opin. Drug Discov. Dev.* 6:353–361.
14. Ballesteros, J., and K. Palczewski. 2001. G-protein-coupled receptors drug discovery: implications from the crystal structure of rhodopsin. *Curr. Opin. Drug Discov. Dev.* 4:561–574.
15. Filipek, S., D. C. Teller, K. Palczewski, and R. Stenkamp. 2003. The crystallographic model of rhodopsin and its use in studies of other G-protein-coupled receptors. *Annu. Rev. Biophys. Biomol. Struct.* 32: 375–397.
16. Shacham, S., M. Topf, N. Avisar, F. Glazer, Y. Marantz, S. Bar-Haim, S. Noiman, Z. Naor, and O. M. Becker. 2001. Modeling the 3D structure of GPCRs from sequence. *Med. Res. Rev.* 21:472–483.
17. Archer-Lahlou, E., I. Tikhonova, C. Escrieut, M. Dufresne, C. Seva, P. Clerc, L. Pradayrol, L. Moroder, B. Maigret, and D. Fourmy. 2005. Modeled structure of a G-protein-coupled receptor: the cholecystokinin-1 receptor. *J. Med. Chem.* 48:180–191.
18. Wank, S. A. 1998. G-protein-coupled receptors in gastrointestinal physiology. I. CCK receptors: an exemplary family. *Am. J. Physiol.* 274:G607–G613.
19. Kennedy, K., C. Escrieut, P. Dufresne, N. Clerc, N. Vaysse, and D. Fourmy. 1995. Identification of a region of the N-terminal of the human CCKA receptor essential for the high affinity interaction with agonist CCK. *Biochem. Biophys. Res. Commun.* 213:845–852.
20. Kennedy, K., V. Gigoux, C. Escrieut, B. Maigret, J. Martinez, L. Moroder, D. Fréhel, D. Gully, N. Vaysse, and D. Fourmy. 1997. Identification of two amino acids of the human cholecystokinin-A receptor that interact with the N-terminal moiety of cholecystokinin. *J. Biol. Chem.* 272:2920–2926.
21. Gigoux, V., C. Escrieut, S. Silvente-Poirot, B. Maigret, L. Gouilleux, J. A. Fehrentz, D. Gully, L. Moroder, N. Vaysse, and D. Fourmy. 1998. Met-195 of the cholecystokinin-A interacts with the sulfated tyrosine of cholecystokinin and is crucial for receptor transition to high affinity state. *J. Biol. Chem.* 273:14380–14386.
22. Gigoux, V., C. Escrieut, J. A. Fehrentz, S. Poirot, B. Maigret, L. Moroder, D. Gully, J. Martinez, N. Vaysse, and D. Fourmy. 1999. Arginine 336 and Asparagine 333 of the human cholecystokinin-A receptor binding site interact with the penultimate aspartic acid and the C-terminal amide of cholecystokinin. *J. Biol. Chem.* 274:20457–20464.
23. Gigoux, V., B. Maigret, C. Escrieut, S. Silvente-Poirot, M. Bouisson, J. A. Fehrentz, L. Moroder, D. Gully, J. Martinez, N. Vaysse, and D. Fourmy. 1999. Arginine 197 of the cholecystokinin-A receptor binding site interacts with the sulfate of the peptide agonist cholecystokinin. *Protein Sci.* 8:2347–2354.
24. Escrieut, C., V. Gigoux, E. Archer, S. Verrier, B. Maigret, R. Behrendt, L. Moroder, E. Bignon, S. Silvente-Poirot, L. Pradayrol, and D. Fourmy. 2002. The biologically crucial C-terminus of cholecystokinin and the non-peptide agonist SR-146,131 share a common binding site in the human CCK1 receptor. *J. Biol. Chem.* 277:7546–7555.
25. Becker, O. M., Y. Marantz, S. Shacham, B. Inbal, A. Heifetz, O. Kalid, S. Bar-Haim, D. Warschaviak, M. Fichman, and S. Noiman. 2004. G-protein-coupled receptors: in silico drug discovery in 3D. *Proc. Natl. Acad. Sci. USA.* 101:11304–11309.
26. Evers, A., and T. Klabunde. 2005. Structure-based drug discovery using GPCR homology modeling: successful virtual screening for antagonists of the α1A adrenergic receptor. *J. Med. Chem.* 48:1088–1097.
27. Huber, T., A. V. Botelho, K. Beyer, and M. F. Brown. 2004. Membrane model for the G-protein-coupled receptor rhodopsin: hydrophobic interface and dynamical structure. *Biophys. J.* 86:2078–2100.
28. Trent, J. O., Z. X. Wang, J. L. Murray, W. Shao, H. Tamamura, N. Fujii, and S. C. Peiper. 2003. Lipid bilayer simulations of CXCR4 with inverse agonists and weak partial agonists. *J. Biol. Chem.* 278:47136–47144.
29. Zhang, Y., Y. Y. Sham, R. Rajamani, J. Gao, and P. S. Portoghese. 2005. Homology modeling and molecular dynamics simulations of the μ-opioid receptor in a membrane-aqueous system. *Chem. Bio. Chem.* 6:1–7.
30. Aburi, M., and P. E. Smith. 2004. Modeling and simulation of the human δ-opioid receptor. *Protein Sci.* 13:1997–2008.
31. Rodinger, T., and R. Pomés. 2005. Enhancing the accuracy, the efficiency and the scope of free energy simulations. *Curr. Opin. Struct. Biol.* 15:164–170.
32. Kalé, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999. NAMD2: greater scalability for parallel molecular dynamics. *J. Comput. Phys.* 151:283–312.
33. Bhandarkar, M., R. Brunner, C. Chipot, A. Dalke, S. Dixit, P. Grayson, J. Gullingsrud, A. Gursoy, W. Humphrey, D. Hurwitz, N. Krawetz, M. Nelson, J. Phillips, A. Shinozaki, G. Zheng, and F. Zhu. 2003. NAMD Users Guide, Vers. 2.5. Theoretical Biophysics Group, University of Illinois and Beckman Institute, Urbana, IL.
34. MacKerell, A. D., Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.

35. Zwanzig, R. W. 1954. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* 22:1420–1426.

36. Gao, J., K. Kuczera, B. Tidor, and M. Karplus. 1989. Hidden thermodynamics of mutant proteins: a molecular dynamics analysis. *Science.* 244:1069–1072.

37. Donnini, S., A. E. Mark, A. H. Juffer, and A. Villa. 2005. Incorporating the effect of ionic strength in free energy calculations using explicit ions. *J. Comput. Chem.* 26:115–122.

38. Straatsma, T. P., H. J. C. Berendsen, and A. J. Stam. 1986. Estimation of statistical errors in molecular simulation calculations. *Mol. Phys.* 57:89–95.

39. Moroder, L., L. Wilschowitz, M. Gemeiner, W. Göhring, S. Knof, R. Scharf, P. Thamm, J. D. Gardner, T. E. Solomon, and E. Wünsch. 1981. Zur Synthese von Cholecystokinin-Pankreozymin. Darstellung von [28-Threonin, 31-Norleucin]-und [28-Threonin, 31-Leucin]-Cholecystokinin-Pankreozymin-(25–33)-Nonapeptid. (Cholecystokinin-pancreozymin synthesis. Synthesis of [28-threonine, 31-norleucine]- and [28=threonine, 31-leucine]cholecystokinin-pancreozymin-(25-33)-nonapeptide.) [in German]. *Z. Physiol. Chem.* 362:929–942.

40. Martinez, J., M. Rodriguez, J. P. Bali, and J. Laur. 1986. Phenethyl ester derivative analogues of the C-terminal tetrapeptide of gastrin as potent gastrin antagonists. *J. Med. Chem.* 29:2201–2206.

41. Fourmy, D., P. Lopez, S. Poirot, J. Jimenez, M. Dufresne, L. Moroder, S. P. Powers, and N. Vaysse. 1989. A new probe for affinity labelling pancreatic cholecystokinin receptor with minor modification of its structure. *Eur. J. Biochem.* 185:397–403.

42. Mehler, E. L., X. Periole, S. A. Hassan, and H. Weinstein. 2002. Key issues in the computational simulation of GPCR function: representation of loop domains. *J. Comput. Aided Mol. Des.* 16:841–853.

43. Dawson, E. S., R. M. Henne, L. J. Miller, and T. P. Lybrand. 2002. Molecular models for cholecystokinin-A receptor. *Pharmacol. Toxicol.* 91:290–296.

44. Giragossian, C., and D. F. Mierke. 2001. Intermolecular interactions between cholecystokinin-8 and the third extracellular loop of the cholecystokinin A receptor. *Biochemistry.* 40:3804–3809.

45. Giragossian, C., and D. F. Mierke. 2003. Determination of ligand-receptor interactions of cholecystokinin by nuclear magnetic resonance. *Life Sci.* 73:705–713.

46. Arlander, S. J., M. Dong, X. Q. Ding, D. I. Pinon, and L. J. Miller. 2004. Key differences in molecular complexes of the cholecystokinin receptor with structurally related peptide agonist, partial agonist, and antagonist. Focus on importance of sulfation of tyrosine in peptide position 27. *Mol. Pharmacol.* 66:545–552.

47. Ding, X. Q., D. I. Pinon, K. E. Furse, T. P. Lybrand, and L. J. Miller. 2002. Refinement of the conformation of a critical region of charge-charge interaction between cholecystokinin and its receptor. *Mol. Pharmacol.* 61:1041–1052.

48. Fourmy, D., C. Escrieut, E. Archer, C. Galés, V. Gigoux, B. Maigret, L. Moroder, S. Silvente-Poirot, J. Martinez, J. A. Fehrentz, and L. Pradayrol. 2002. Structure of cholecystokinin receptor binding sites and mechanism of activation/inactivation by agonists/antagonists. *Mol. Pharmacol.* 91:313–320.

49. Lanyi, J. K., and H. Luecke. 2001. Bacteriorhodopsin. *Curr. Opin. Struct. Biol.* 11:415–419.

50. Lu, N., D. A. Kofke, and T. B. Woolf. 2004. Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods. *J. Comput. Chem.* 25:28–39.

51. Bennett, C. H. 1976. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* 22:245–268.

52. Hibert, M. F., S. Trumpp-Kallmeyer, J. Hoflack, and A. Bruinvels. 1993. This is not a G-protein-coupled receptor. *Trends Pharmacol. Sci.* 14:7–12.

53. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD—visual molecular dynamics. *J. Mol. Graph.* 14:33–38.